# Input frequency and talker variability affect child word production

Jessica Filson

Oklahoma State University

Fall 2018

## Abstract

**Purpose:** Many factors affect how children learn to produce new words, one being variability. Some recent research suggests that variability is beneficial in learning to produce new words, while some suggests it is not beneficial. This study aims to determine the effect of input frequency and talker variability on how children learn to produce new words.

**Method:** Thirty-six typically developing preschool-age children participated in a novel word production task. The participants were first introduced to the nonwords in a pretest, where they heard the words one time before having a chance to make a production. In the test block, participants heard nonwords 1, 3, 6, or 10 times by either by a single talker or by multiple talkers before producing the word.

**Results:** Productions were found to be more accurate in the test phase than during the pretest. Talker variability was found to have no significant effect during either the pretest or test phases. In the pretest block, input frequency did not have a significant effect, but a significant effect of input frequency was found during the test block. Input frequencies of 3, 6, and 10 resulted in more accurate productions than an input frequency of 1; however, there was no significant difference between input frequencies of 3, 6, or 10.

**Conclusions:** The results suggest that variability is beneficial to some extent when learning to produce new words, and more exposures to a word can lead to more accurate productions as implied by more production accuracy during the test phase. These results can be used in a clinical setting by Speech-Language Pathologists when adjusting therapy frequency for specific clients. The lack of significant differences between input frequencies of 3, 6, and 10 suggests that different clients could benefit from varying frequencies of therapy, which gives clinicians the freedom to determine what is most beneficial for individual clients.

# Input frequency and talker variability affect child word production

When looking at how children learn to produce new words, many factors play a role. Producing new words is a complex process and can be affected by many influences. One factor that can affect accuracy of new word production is variability. In these cases, variability can refer to input frequency or talker variability. Input frequency refers to how often children hear a target word (Richtsmeier and Good, 2018), and talker variability refers to how many different speakers produce the target word for the child. Much of the previous research shows that variability in what kids hear seems to be beneficial, although that is not always found to be the case. With this in mind, the current experiment looks at to what extent input frequency and talker variability are beneficial in how children produce new words. To begin the study, we must begin with an overview of what previous research has found concerning the effects of variability on new word production. Since variability is a very general term and, as aforementioned, it includes multiple types of variability, a range of effects of variability on learning to produce new words have been found. In the following sections, we will look at the positive and negative effects of variability, as well as the effects of producing or listening to a word first when learning to accurately produce new words.

## Variability is Beneficial

In their 2011 study, Plante et al. discuss whether or not “the frequency of speech input to children with specific language impairment (SLI) and their typically developing peers is sufficient to effect word production” (Plante et al. 2011). Specific language impairment refers to a language impairment in the context of normal intelligence. Usually, children produce sounds that are more frequent in their native language, so this experiment was designed to see if that could be demonstrated in a short time period in both children with SLI and typically developing (TD) children. This experiment included 64 native English speakers—32 with SLI and 32 with TD. The materials included 8 CVCCVC nonwords like /fæmpəm/ and /mæstəm/. Each child was assigned two nonwords with a medial consonant cluster frequently found in English and two with a medial consonant cluster infrequently found in English. The nonwords were all paired with pictures of cartoon animals. The test included a familiarization phase and a test phase. During the familiarization phase, the children heard the nonwords with high experimental frequency sounds 10 times and the nonwords with low experimental frequency sounds one time; they were also shown a picture associated with each nonword. During the test phase the children produced each of the four nonwords four times. The test considered both production accuracy as well as the time taken to respond, which was calculated for all responses. The results showed that overall the SLI group was less accurate and took more time to respond than did the TD group. All children were less accurate in their productions of words including low frequency English sounds as well as words with a lower experimental frequency. Both groups also responded to nonwords with a high English frequency quicker than those with low English frequency, and they both responded quicker to the nonwords presented more frequently during the experiment. The researchers found that experiment frequency can affect accuracy of word production in a short amount of time. They also concluded that talker variability increased the response speed and accuracy for both groups.

As stated in the aforementioned article, talker variability and increased input frequency led to positive effects on word learning for both typically developing children and those with SLI. Similarly, Aguilar, Plante, and Sandoval (2018) look at how object variability affects the semantic learning of new words by preschoolers with SLI. The researchers claim that lexical variation is shown to help those with language impairment learn words because “variation around the target is thought to make the target of learning perceptually salient and therefore easier to learn” (3). Similar to previous research, the researchers designed a study to see whether or not object variability facilitates the semantic learning of new words in children with SLI. They hypothesized that children with SLI in a high exemplar variability condition will learn more words within the experiment than children not exposed to any exemplar variability. The participants included 18 4- and 5-year-old children with SLI. Each was administered tests to confirm their diagnosis of SLI. They were split into two equal groups: the no variability group in which there were 3 identical objects corresponding to each new word during training and the high variability group in which there were 3 different objects corresponding to each new word during training. The target words included 8 1- and 2-syllable nouns, such as “burlap” and “sequin”, which were deemed unfamiliar to the children. The words were taught to the children during a 3 day training during which the words were presented within 3 activities. The participants were tested on their ability to remember the words 1 day after training and again 3 weeks after training. The vocabulary tests were administered by 4 research clinicians and they included both objects seen during training and unfamiliar objects. The test was in the format of a flip book where the children had to point to the correct picture corresponding to a word. Results showed that there was no distinction between the no variability and high variability groups in their word learning on the test 1 day after training. However, after 3 weeks, the high variability group showed a greater retention for the new words than did the no variability group. These findings support the researchers’ hypothesis and the previous data that shows the benefit of object variability on word learning.

**Variability is Not Always Beneficial**

The results of the two previous studies show the positive effect of variability on word learning in both typical developing and language-impaired children. In these cases, variability occurred in both what the kids heard as well as in the object corresponding to each word. However, not all research supports the claim that variability is beneficial in word-learning. For example, in "Frequencies in perception and production differentially affect child speech," Richtsmeier and Good (2018) look at how varying input frequencies, or experimental frequencies as referred to in the Plante et al. study, affect production accuracy and referent identification. Previous research generally shows that perceptual and production frequencies benefit speech production and word learning when analyzed independently, but when combined, the effects are not always simple. Richtsmeier and Good's study was designed to answer the question "is more always better?" when it comes to input frequency and how it affects speech accuracy and referent identification. In order to test this, they exposed 41 typically developing children between the ages of 3;0 and 4;8 (years; months) to nonwords, such as /mæfpəg/ and /fugdən/, with input frequencies of 1, 3, 6, and 10. They also presented the words with two talker variability conditions: single talker or multiple talkers. The experiment consisted of the test, where the input frequencies were manipulated and the children had 3 chances to say each nonword, ABX discrimination (discriminating between minimal pairs), referent identification (identifying the make-believe animal picture when given the corresponding nonword), a posttest, and a nonword repetition task. In the posttest, it did not appear that children were affected by input frequency. In conclusion, the results of this experiment show that more input is not always better for production accuracy.

**Listening and Production**

In many studies that focus on the effect of variability on how children learn words, the participants were expected to produce the new words, such as in Richtsmeier and Good's, as well as Plante et al.'s, study. Word production is often thought to be a useful method in learning new words. However, some research shows that production can actually be detrimental on word-learning. This concept is tested in "Reverse production effect: Children recognize novel words better when they are heard rather than produced". Zamuner et al. (2017) examine to what extent speech production affects lexical acquisition, as measured by novel word recognition. Much of the previous research shows that production is beneficial in word-learning, rather than just hearing the new word. However, some studies, such as Schwartz and Leonard's 1982 study, claim that children will not necessarily benefit from production if the words are outside of their production repertoire. Zamuner et al.'s current study was designed to provide "controlled data on the effects of production on word learning" in children. This experiment contained two parts: Experiment 1 and Experiment 2. In the first experiment, participants were shown a picture of a make-believe animal and the associated nonword was played. They were instructed to either listen to or produce the word. The experiment used an eye gaze tracking task to test whether the children looked longer at the words they heard or produced. Results showed that the children looked at the targets of the words they heard more than the ones they produced, but there was no statistically significant difference in the slope or shape of the looking curves for each condition. Experiment 2 followed the same procedure, but this one included a recognition task at the end of the experiment. The participants were asked if they remembered the names of any of the animals; results showed that the participants were able to accurately map targets for both the words they heard and produced, but they looked more at the targets they heard, and there was a greater recall of the heard targets. In this case, it seems to matter greatly whether the child listens to or produces the word first when learning new words. Zamuner et al. suggest that these findings may differ from much of the previous research due to the fact that the children's language skills are not yet fully developed, so they might not benefit from producing the unknown words in a difficult task.

## Current Study

Similar to these studies, the current study looks at factors that could play a role in how preschool-aged children learn to produce new words. Using knowledge of the previous research, we manipulated the variability of these factors that affect production accuracy and referent identification of new words. Specifically, the factors we manipulated are input frequency and talker variability. These variables are the same as those used in Richtsmeier and Good's study; however, in the current study, listening and production were reversed in order to test whether or not the order of production mattered when it comes to how children learn to accurately produce new words. Like Richtsmeier and Good's study, the current study uses a structure that allows the participants to produce the word immediately after exposure. This structure is used in order to more accurately determine whether or not input frequency and talker variability had a direct effect on the child's productions, rather than exposing the child to the word at the beginning of the study and relying on his or her memory to accurately produce the word.

# Method

## Participants

Thirty-six children between the ages of 3;2 and 5;1 were recruited for this study. They were recruited through local daycares and preschools, as well as through advertisements on Facebook. The children's parents then contacted the researchers in order to set up a time for the children to come in for the experiment. Before beginning the experiment, the parents signed a consent form, and each child marked an assent form after the experimenter explained the procedure of the study. All of the participants met criteria for typical development. Each child completed a pure tone hearing screening and had to respond to tones of 1000, 2000, and 4000 Hz at 25 dB SPL. The participants' parents also filled out a parent questionnaire in order to make sure the child met the criteria for typical development. The participants whose responses were used in the data analysis each received a standard score of 85 or above on the Goldman-Fristoe Test of Articulation-2 (GFTA-2), which met the criteria for typical development. The average standard score was 108.6, which indicates that most of the participants fell in the high average range. The researchers reviewed each participant's developmental history using the questionnaire completed by the parents. Every child was reported as typically developing, and only one participant had a history of speech therapy. Additional measures were collected to provide a fuller representation of each participant's speech and language; these measures included an auditory discrimination task, as well as a nonword repetition task, which was based on Dollaghan and Campbell (1998). However, these measures are not reported here. One child's score on the GFTA-2 did not meet the criteria for typical development, and one child's responses were not properly recorded due to experimenter error, so their responses are not included in the data analysis. The remaining 34 participants' responses were included in the data analysis.

**Materials**

The materials included eight CVCCVC nonwords. These included /pɛmtəs/, /niʃkət/, /mæfpəg/, /fugdən/, /sabləf/, /tʌvʧəp/, /bozjəm/, and /gɪsnək/. Each nonword contained a word-medial consonant sequence, and these sequences were not included in the other words. For example, in the nonword /bozjəm/, the word-medial /z/ and /j/ do not appear in the medial position in any of the other nonwords used. The nonwords differed in input frequency conditions, where the number of exposures prior to production varied. The words were placed in four different frequency conditions; two were heard 1 time, two were heard 3 times, two were heard 6 times, and 2 were heard 10 times. The nonwords were also divided into two conditions of talker variability, where either a single talker or multiple talkers presented each nonword. When the nonword was presented by multiple talkers, the number of talkers corresponded to the number of input frequencies. For example, if /sabləf/ had an input frequency of 10, the nonword would be presented by 10 different talkers. The nonwords with an input frequency of 1 were consequently produced by a single talker.

In addition, a hand-drawn picture of a make believe animal corresponded to each nonword, in order to provide a visual representation of the nonwords. Eight make-believe animals were used to represent the eight nonwords. Since each animal corresponded with one nonword, it is unlikely the pictures would skew the results of the study regardless of the input frequency of talker variability.

**Procedure**

The experiment consisted of five “blocks”, or sections of the experiment with different purposes. The blocks included pretest, auditory discrimination, referent identification, test, and a nonword repetition task based on Dollaghan and Campbell (1998). The test and pretest consisted of trials, or opportunities to produce a nonword. During the pretest, there were 2 trials per nonword, and during the test, there were three trials per nonword. In the test block, each trial contained a various number of “exposures” of the nonwords. The trials could include either 1, 3, 6, or 10 exposures to the nonword before the child was given an opportunity to produce the nonword.

The children were brought to the research lab for the experiment by their parents. The parents observed in the experiment room, but they did not help the children with the experiment. They were able, however, to provide encouragement if the children became restless or upset. The children sat at a child-size table which had a computer in the center with speakers on both sides and a mouse in front. The speakers were set at a comfortable volume and was consistent across every participant. The experiment was controlled by the experimenter using Paradigm computer software (Paradigm, 2016). The experimenter sat on the left side of the child and presented the directions.

The children were familiarized with the structure of the experiment during a training phase using real English words—*ball, kitty,* and *donut.* The children completed four training trials (the trial with *ball* was repeated), and the experimenter helped the children understand the experiment directions. The experimenter explained to the child that they would be playing a game in which they heard a word and had to repeat it after a set number of exposures. After completing the training trials, the experimenter explained that the child would hear the names of silly make-believe animals. At the top left of the screen was a picture of the make-believe animal. During the pretest phase, the experiment screen consisted of blue boxes and a yellow box as shown in Figure 1. The children were told to listen to the word when the box was blue and say the word when the box turned yellow. The make-believe animal moved above the next box with each exposure. Each nonword was presented 1 time prior to the child's chance to produce the word. Once the animal got to the yellow box, the child was to produce the nonword. If he or she did not produce the nonword, the experimenter would prompt the child saying "what is this animal's name?" If no response was given, the experimenter would continue to the next trial. The children were able to control the pace of the nonword exposures by touching the screen after each exposure. If the child did not wish to do this, the experimenter could control the pace by clicking the mouse as well. The experimenter also controlled the experiment when a new trial began.

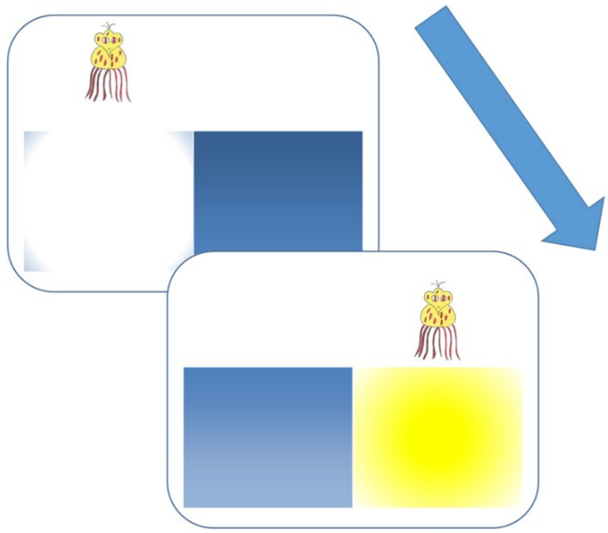

Figure 1. This figure shows a pretest trial with an exposure of 1.

The second block of the experiment was an auditory discrimination task and the third block of the experiment was a referent identification task, neither of which is discussed here. The fourth block was the test phase. During this block, each child was given either 1, 3, 6, or 10 exposures of each nonword per trial. The test phase followed the same structure of the pretest, as shown in Figure 2, but the make-believe animals were presented in 3 trials each.

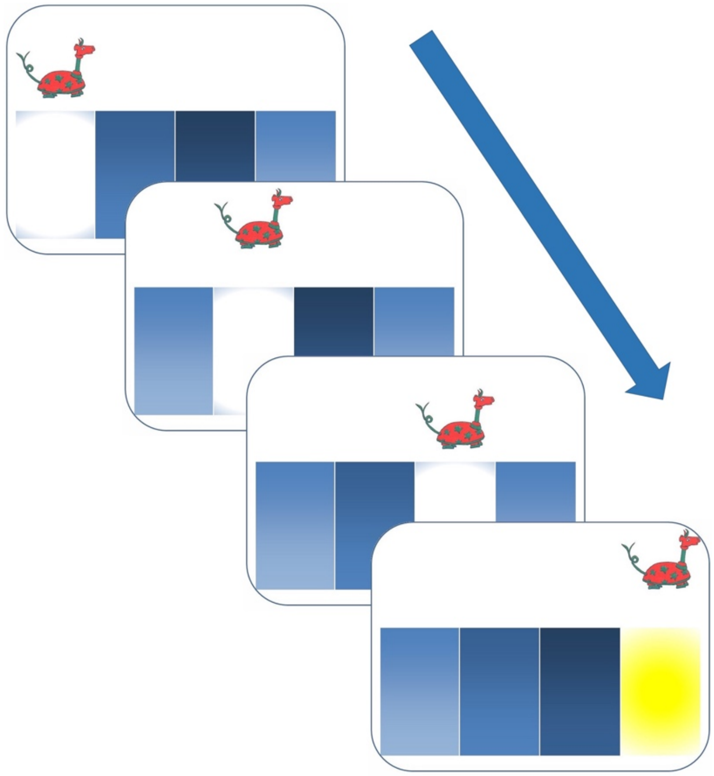

Figure 2. This figure shows a trial from the test block with an exposure of 3. The child was told to listen when the animal was above the blue boxes and speak when the animal arrived at the yellow block.

After completing the test during the fourth block, the fifth block of the experiment was a nonword repetition task using 16 nonwords from Dollaghan and Campbell (1998). Sticker sheets were used as a reinforcement between each block of the experiment, and after the entire experiment was completed, each child was able to choose a prize from the treasure chest. The study typically took 40-60 minutes to complete and families were compensated $20 for bringing their child in for the experiment.

**Analysis**

The children's productions were transcribed phonetically by the research assistant. The transcriptions of the four consonants in each nonword were scored on a 3-point scale, where a score of 3 suggested a correct production of the target consonant in terms of manner of articulation, place or articulation, and voicing. A score of 2 indicated that two of the three features were produced correctly. A score of 1 indicated that an attempt at production was made, but there were two or more features that were produced incorrectly. A score of 0 corresponded to a deletion of that consonant. Each transcription could receive a score up to 12 possible points.

For example, if the target nonword was /niʃkət/ and the participant said /niskI/, the child would receive a score of 3 for the initial consonant /n/, a score of 2 for the second consonant since the produced consonant and the target consonant agreed in terms of manner and voicing, 3 points for a correct production of /k/, and 0 points for the deletion of the final consonant /t/. After totaling these scores, the child would receive a total score of 8 out of 12 for his or her production.

## Results

The data was analyzed using an analysis of variance (ANOVA). Since talker variability only occurs in input frequencies of 3, 6, and 10, the scores for production accuracy were examined in a 2×2×3 design (Test [pretest, test] × Talker [single talker, multiple talkers] × Input Frequency [3, 6, 10]). A significant effect was found for pretest-test, $p < .001$, with productions made during the test phase being more accurate, as shown in Figure 3. There was no significant effect of talker variability, nor was there a significant effect of the three input frequencies tested in this analysis.

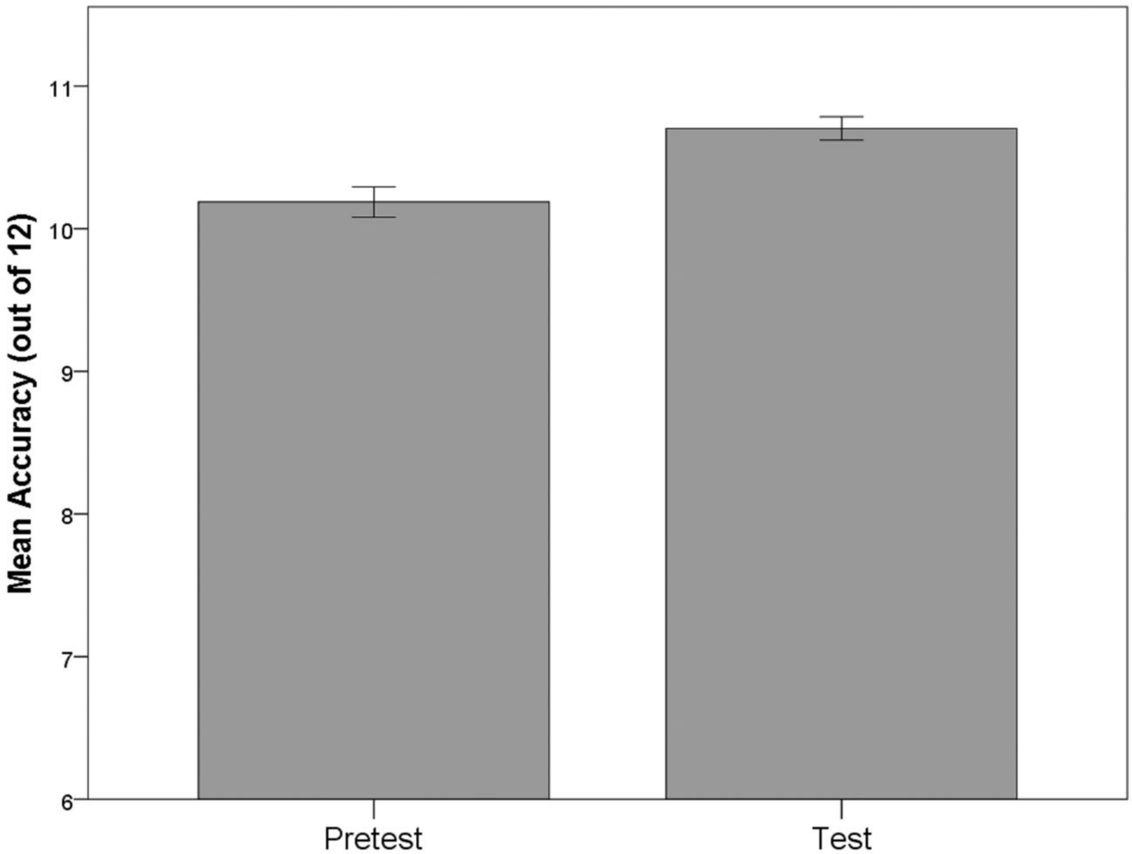


Figure 3. Mean word accuracy out of 12 (y-axis) by pretest-test (x-axis).

Only input frequencies of 3, 6, and 10 were examined in the first analysis. Similar to the study by Richtsmeier and Good (2018), we conducted a second ANOVA to look at the effects of all four input frequencies in a 2×4 design (Test [pretest, test] × Input Frequency [1, 3, 6, 10]. A significant effect of pretest-test was once again found, p=.001, with participants having more accurate productions in the test phase. A significant effect was also found in the interaction between Test and input frequency, p=.012.

To break down the data even more, the effects of input frequency (1, 3, 6, 10) on both pretest and test were looked at individually. There was no significant effect of input frequency on the mean production accuracy in the pretest. In the test, however, input frequency had a significant effect on accuracy, p=.001.

From here, the effect on accuracy of each input frequency was examined. Significant differences in accuracy between input frequencies of 3, 6, and 10 were not found, but these three input frequencies resulted in more accurate productions than an input frequency of 1 in the test phase, p=.011. Figure 4 shows the mean accuracy of all four input frequencies [1, 3, 6, 10] in both the pretest and test.

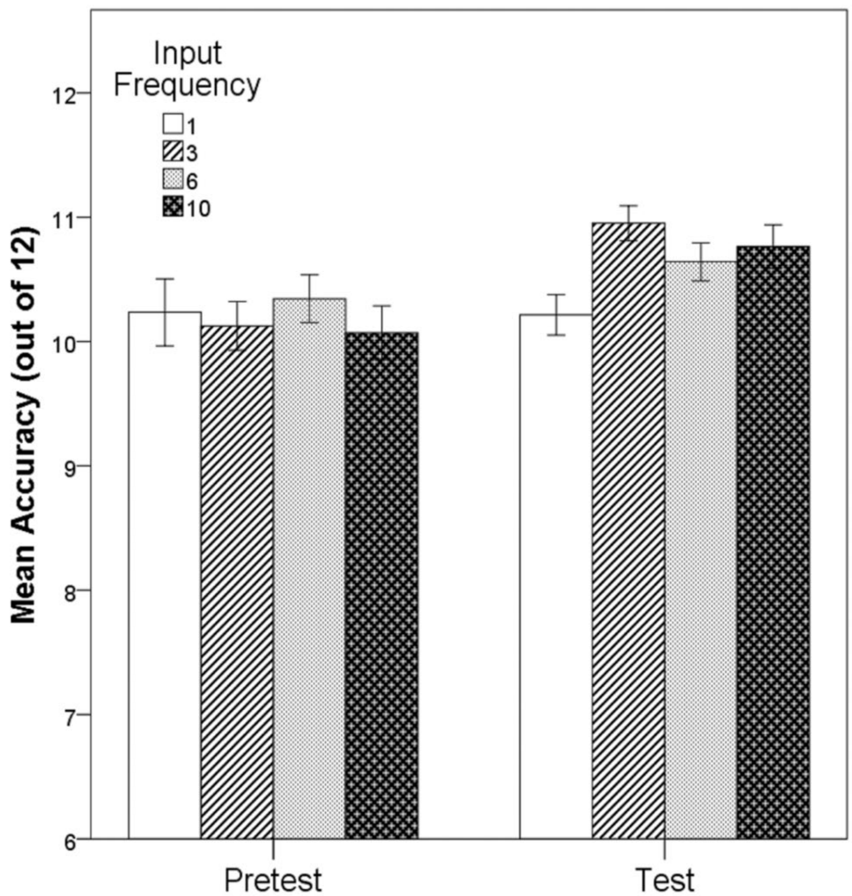


Figure 4. Mean word accuracy out of 12 (y-axis) by pretest-test (x-axis) and input frequencies (represented by the vertical bars).

## General Discussion

In the current study, children were examined on their ability to produce nonwords in the form of make-believe animals. The children were familiarized with the nonwords by either single or multiple talkers and with four different input frequencies—1, 3, 6, and 10. The children produced the nonwords in a pretest and test block.

This study's use of a pretest-test design gives us insight into the effects of the order of listening and production on learning new words. Opposite of Richtsmeier and Good's study (2018), the current study contained a pretest during which each participant produced each nonword after hearing it only one time. The test phase gave children the opportunity to produce each word three times. With the test phase taking place after the pretest, child had more exposures to each nonword, resulting in higher production accuracy during the test phase.

Results show that overall children produced the nonwords more accurately in the test than in the pretest. This is to be expected since the children were only exposed to the words one time each during the pretest. Talker variability was not found to have a significant effect on production accuracy in either the pretest or test blocks. These findings contrast with Plante et al.'s study (2011) that claims talker variability has a positive effect on production accuracy in both typically developing children and children with specific language impairment, although the effect of talker variability was in the same direction as in Plante et al.'s study, with a numerical benefit for hearing multiple talkers. The difference between Plante et al.'s study and this one could be due to the level of difficulty associated with the nonwords and the participants' abilities to produce the consonants of each nonword accurately. However, because we used a within-subjects study design, like Richtsmeier and Good (2018), our findings on the lack of significance of talker variability are valuable.

The results of the current study also show that input frequency did affect accuracy in the test phase. Within the test phase, an input frequency of 1 resulted in less accurate productions than when the input frequency was 3, 6, or 10. A significant difference between input frequencies of 3, 6, and 10 was not found, but collectively there was a significant difference when compared to an input frequency of 1.

In the introduction, we looked at studies that reported variability is both beneficial (Aguilar, Plante, & Sandoval, 2018; Plante et al., 2011) and not always beneficial (Richtsmeier and Good, 2018) in word learning. With that being said, the current study arguably combines the results of these previously mentioned studies. In terms of talker variability, our results show that while variation of talkers is not necessarily hurtful in learning new words, it is not helpful either. Input variation, however, is found to be beneficial in learning new words. As aforementioned, our results did not specify which input frequency (1, 3, 6, 10) was most helpful in accurate productions, but input frequencies above 1 (3, 6, 10) were found to be significantly more beneficial than only one exposure to the nonword.

These results may be valuable for Speech-Language Pathologists (SLP) in a clinical setting. Depending on the client's goals, increased frequency of therapy or longer sessions may result in higher achievement of goals. However, like Richtsmeier and Good's study (2018) that found that more is not always better in terms of exposures when learning new words, more therapy might not always be better in helping clients achieve goals. Since our study did not find a significant difference in production accuracy between being exposed to the word either 3, 6, or 10 times, clinicians can use this study as a basis from which to manipulate therapy to determine what works best for their individual clients.

## Conclusion

This study provides a valuable basis from which to conduct further research. This study involved preschool-age children, like many of the previous studies involving the production of new words. With the participants being so young, at times it was difficult for them to focus and stay motivated throughout the experiment. Future studies could benefit from involving participants from a different age group in order to determine if the same effects of variability on production accuracy occur. Unlike some of the previous studies (Aguilar, Plante, & Sandoval, 2018; Plante et al., 2011), this study only involved typically developing children. Future research could include children with language impairments in order to determine if the same effects of input frequency and talker variability are found. Additionally, since no significant effect was found between input frequencies of 3, 6, and 10, future studies could adjust input frequencies to give the participants either more or fewer exposures in order to determine if there is a specific input frequency that elicits the most accuracy in new word production.

## References

Aguilar, J., Plante, E., & Sandoval, M. (2018). Exemplar Variability Facilitates Retention of Word Learning by Children With Specific Language Impairment. *Language Speech And Hearing Services In Schools*, *49*(1), 72. doi: 10.1044/2017_lshss-17-0031

Plante, E., Bahl, M., Vance, R., & Gerken, L. (2011). Beyond phonotactic frequency: Presentation frequency effects word productions in specific language impairment. *Journal Of Communication Disorders*, *44*(1), 91-102. doi: 10.1016/j.jcomdis.2010.07.005

Richtsmeier, P., & Good, A. (2018). Frequencies in perception and production differentially affect child speech. *Journal of Speech, Language, and Hearing Research,* 1-15. doi: https://jslhr.pubs.asha.org/article.aspx?articleid=2718185

Zamuner, T., Strahm, S., Morin-Lessard, E., & Page, M. (2017). Reverse production effect: children recognize novel words better when they are heard rather than produced. *Developmental Science*, *21*(4), e12636. doi: 10.1111/desc.12636